**Crop Recommendation and Yield Production:**

--------------------------------------------------------------------------------------

Aim:

1. Achieving maximum crop yield at minimum cost is one of the goals of agricultural production.
2. Early detection and management of problems associated with crop yield indicators can help increase yield and subsequent profit.
3. Predictions could be used by crop managers to minimize losses when unfavourable conditions may occur.
4. Climate and other environmental changes in the world and Indian sub-continent has become a major threat to the agricultural economy.
5. Being able to predict future yield of specific crop would help farmers, decide the better option among many to maximize the yield.
6. Yield production can reduce the financial risk of farmers.

Dataset:

1. Crop\_ph.csv
2. CropData.csv
3. CropRequiredTemperature.csv
4. District\_ph.csv
5. MaharashtrastateRainfall.csv

Algorithm:

We have divided our project in three parts:

**1. Area Production:**

We used Cropdata.csv dataset here. It consists of State name, District, Crop name, year, Season, Crop Area, and Production. It has 12629 entries of such data of all the districts in the state of Maharastra.

Lists of 34 prominant cities are shown to choose from all the districts.

User is asked to enter the number corresponding to the city.

The data-set CropData.csv is loaded in “areaproductiondata” variable.

Then the entries corresponding to chosen city is loaded. This is then printed for confirmation of loading process.

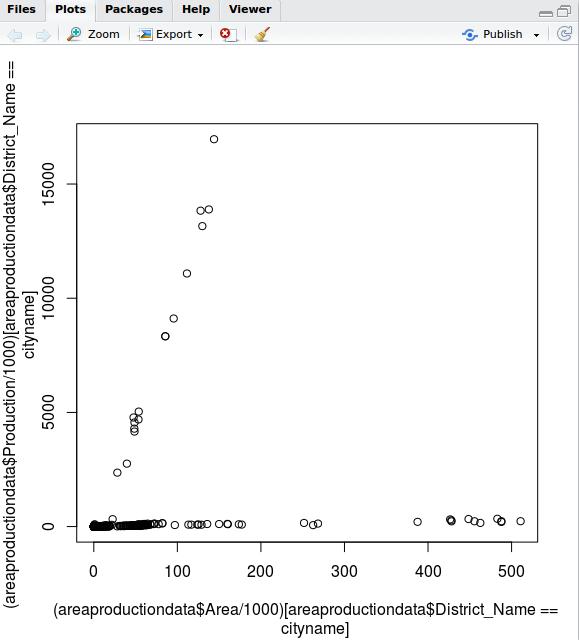
Then a plot of area vs. production is plotted.(Both the values are divided by 1000 for better visualization purposes).

Maximum area/production value is calculated and the corresponding entry details are printed

Average area/production is calculated and printed.

Then all above average crops (area/production) are segregated from the areaproductiondata, and a list of useful crops are printed.

Similarly all the crops below average are then printed as un useful crops



1. **Temperature Rain Corelation.**

We used MaharashtrastateRainfall.csv dataset here. It consists ofState name**,** District name, Year**.** All the 12 months from Jan-December Annual Total rainfall.

We used TempRainfallData.csv dataset here. It consists of City, Month, Year, Period, Mean Temperature, Mean Rainfall.

We used CropRequiredTemperature.csv dataset here. It consists of Crop name, Min temperature, Max temperature, Min rainfall, Max rainfall. It has 12 such entries.

Lists of 34 prominent cities are shown to choose from all the districts.

User is asked to enter the number corresponding to the city. The data-set is loaded in ‘rainfallavailbaledata’ variable.

Few of the starting entries are shown to confirm that he data is loaded.

The entries corresponding to the selects city are loaded as a vector and mean is calculated of the annul\_total.

The temprainfall.csv dataset is then loaded.

A linear model is created between mean temp and mean rainfall

Then croprequiredtemparature dataset is loaded and printed for confirmation.

Then crops having required minimum-maximum rainfall falling inside average rainfall are found out.

1. **pH Contribution**

We used District\_ph.csv dataset here. It consists of City name, pH min ,pH max.

We used Crop\_ph.csv dataset here. It consists of Crop name, pH required, pH min, range

Lists of 34 prominent cities are shown to choose from all the districts.

User is asked to enter the number corresponding to the city.

The data-set District\_ph..csv and Crop\_ph..csv is loaded in ‘data1’ and ‘data2’ variable respectively. Entry corresponding to selected city is loaded and printed.

Crops ph\_req and ph\_max\_range falling under the range of city’s ph\_min and ph\_max are selected.

Max\_pH outliers are plotted

Merged data from data1 and data2 is showed.

This Ouput gives us which Crop is most suitable for that particular city with the given pH values.

Conclusion:

* The analysis of rainfall and temperature areawise helps us determine the most suitable range required for different crops.
* Analyzing pH values of land with the above analysis of rainfall and temperature, a regression model is created to help predict which crop is most suitable in which area.
* Area vs Production helps to reinforce the correctness of this model.

Code:

cat("Select City Number:\n\n1 AHMEDNAGAR\n2 AKOLA\n3 AMRAVATI\n4 AURANGABAD\n5 BEED\n6 B`HANDARA\n7 BULDHANA\n8 CHANDRAPUR\n9 DHULE\n10 GADCHIROLI\n11 GONDIA\n12 HINGOLI\n13 JALGAON\n14 JALNA\n15 KOLHAPUR\n16 LATUR\n17 NAGPUR\n18 NANDED\n19 NANDURBAR\n20 NASHIK\n21 OSMANABAD\n22 PARBHANI\n23 PUNE\n24 RAIGAD\n25 RATNAGIRI\n26 SANGLI\n27 SATARA\n28 SINDHUDURG\n29 SOLAPUR\n30 THANE\n31 WARDHA\n32 WASHIM\n33 YAVATMAL\n34 MUMBAI\n\nChoose Corresponding Number to Your City from the Given List:\n")

readinteger <- function()

{

n <- readline(prompt="Enter City Number: ")

n <- as.integer(n)

if (is.na(n)){

n <- readinteger()

}

return(n)

}

y <- readinteger()

cityname <- switch(y,"AHMEDNAGAR","AKOLA","AMRAVATI","AURANGABAD","BEED","BHANDARA","BULDHANA","CHANDRAPUR","DHULE","GADCHIROLI","GONDIA","HINGOLI","JALGAON","JALNA","KOLHAPUR","LATUR","NAGPUR","NANDED","NANDURBAR","NASHIK","OSMANABAD","PARBHANI","PUNE","RAIGAD","RATNAGIRI","SANGLI","SATARA","SINDHUDURG","SOLAPUR","THANE","WARDHA","WASHIM","YAVATMAL","MUMBAI")

print(cityname)

data1 <- read.csv("C:/Users/rad/Downloads/District\_ph.csv")

data2 <- read.csv("C:/Users/rad/Downloads/Crop\_ph.csv")

subval1 <- subset(data1, data1$City == cityname, select = c("City","pH\_min","pH\_max"))

print(subval1)

result1 <- subset(data2,data2$pH\_Req <= subval1$pH\_min & data2$pH\_max\_Range >= subval1$pH\_max, select = c("Crop"))

par(mfrow=c(2, 3))

boxplot(data1$pH\_max, main="Max\_pH", sub=paste("Outlier rows: ", boxplot.stats(data1$pH\_max)$out))

boxplot(data1$pH\_min, main="Min\_pH", sub=paste("Outlier rows: ", boxplot.stats(data1$pH\_min)$out))

boxplot(data2$pH\_Req, main="Req\_pH", sub=paste("Outlier rows: ", boxplot.stats(data2$pH\_Req)$out))

boxplot(data2$pH\_max\_Range, main="Max\_pH\_Range", sub=paste("Outlier rows: ", boxplot.stats(data2$pH\_max\_Range)$out))

boxplot(data2$pH\_min\_Range, main="Min\_pH\_Range", sub=paste("Outlier rows: ", boxplot.stats(data2$pH\_min\_Range)$out))

library(e1071)

par(mfrow=c(1, 2))

plot(density(data1$pH\_min), main="Density Plot: City\_pH\_min", ylab="Plot", sub=paste("Skewness:", round(e1071::skewness(data1$pH\_min), 2)))

polygon(density(data1$pH\_min), col="red")

plot(density(data2$pH\_Req), main="Density Plot: Crop\_pH\_Req", ylab="Plot", sub=paste("Skewness:", round(e1071::skewness(data2$pH\_Req), 2)))

polygon(density(data2$pH\_Req), col="red")

print(result1)

rainfallavailabledata <- read.csv("C:/Users/rad/Downloads/MaharashtrastateRainfall.csv")

head(rainfallavailabledata)

subvalrainfall <- subset(rainfallavailabledata, rainfallavailabledata$District == cityname, select = c("State","District","Year","January","February","March","April","May","June","July","August","September","October","November","December","Annual\_Total"))

convertednumbers <- as.numeric(as.vector(subvalrainfall$Annual\_Total))

totalrainfall = mean(convertednumbers,na.rm=TRUE)

avgrainfall = totalrainfall/12

print(avgrainfall)

par(mfrow=c(1, 1))

plot(density(as.numeric(subvalrainfall$Annual\_Total)), main="Density Plot: City\_Annual\_Rainfall", ylab="Plot", sub=paste("Skewness:", round(e1071::skewness(as.numeric(subvalrainfall$Annual\_Total)), 2)))

polygon(density(as.numeric(subvalrainfall$Annual\_Total)), col="red")

croprainfalltempdata <- read.csv("C:/Users/rad/Downloads/CropRequiredTemperature.csv")

head(croprainfalltempdata)

result2 <- subset(croprainfalltempdata, (croprainfalltempdata$min\_rainfall<=avgrainfall & croprainfalltempdata$max\_rainfall>=avgrainfall), select = c("Crop"))

par(mfrow=c(1, 2))

plot(density(as.numeric(croprainfalltempdata$max\_rainfall)), main="Density Plot: Crop\_Max\_Rainfall", ylab="Plot", sub=paste("Skewness:", round(e1071::skewness(as.numeric(croprainfalltempdata$max\_rainfall)), 2)))

polygon(density(as.numeric(croprainfalltempdata$max\_rainfall)), col="red")

plot(density(as.numeric(croprainfalltempdata$min\_rainfall)), main="Density Plot: Crop\_Min\_Rainfall", ylab="Plot", sub=paste("Skewness:", round(e1071::skewness(as.numeric(croprainfalltempdata$min\_rainfall)), 2)))

polygon(density(as.numeric(croprainfalltempdata$min\_rainfall)), col="red")

print(result2)

areaproductiondata <- read.csv("C:/Users/rad/Downloads/CropData.csv")

head(areaproductiondata)

subval <- subset(areaproductiondata, areaproductiondata$District\_Name == cityname, select = c("State\_Name","District\_Name","Crop\_Year","Season","Crop","Area","Production"))

print(subval)

par(mfrow=c(1, 1))

plot((areaproductiondata$Area/1000)[areaproductiondata$District\_Name==cityname],(areaproductiondata$Production/1000)[areaproductiondata$District\_Name==cityname],main = "Area vs Production")

plot((areaproductiondata$Area)[areaproductiondata$District\_Name==cityname],(areaproductiondata$Production)[areaproductiondata$District\_Name==cityname],main = "Area vs Production")

abline(lm(subval$Production~subval$Area))

plot((areaproductiondata$Crop\_Year)[areaproductiondata$District\_Name==cityname],(areaproductiondata$Production)[areaproductiondata$District\_Name==cityname],main = "year vs Production")

max((subval$Production)/(subval$Area),na.rm=TRUE)

expectedcropname <- subval[which.max(subval$Production/subval$Area),]

print(expectedcropname)

print(expectedcropname$Crop)

avgcrop<- mean((subval$Production)/(subval$Area),na.rm=TRUE)

print(avgcrop)

result3 <- subset(subval, (subval$Production/subval$Area)>avgcrop, select = c("Crop"))

print(unique(result3))

#x=max((climatecitydata$Production/climatecitydata$Area),na.rm=TRUE,subset=climatecitydata$District\_Name=="YAVATMAL")

#print(x)

unusefulcrop <- subval[which.min(subval$Production/subval$Area),]

print(unusefulcrop)

print(unusefulcrop$Crop)

#install.packages("dplyr")

library(dplyr)

cat("Crop Predicted By Considering PH Value and Rainfall Factors:\t")

print(unique(intersect(as.vector(result2), as.vector(result1))))

cat("Crop Predicted By Considering Rainfall and Area-Production Factors:\t")

print(unique(intersect(as.vector(result2), as.vector(result3))))

cat("Crop Predicted By Considering PH Value and Area-Production Factors:\t")

print(unique(intersect(as.vector(result1), as.vector(result3))))

cat("Crop Predicted By Considering All Three Factors Factors:\t")

rs12<-unique(intersect(as.vector(result2), as.vector(result1)))

rs3<-as.vector(result3)

print(unique(intersect(rs12,rs3)))